# CAPTURING VARIABILITY IN FLIPPED MATHEMATICS INSTRUCTION

Samuel Otten
University of Missouri
ottensa@missouri.edu

Zandra de Araujo
University of Missouri
dearauojz@missouri.edu

Milan Sherman
Drake University
milan.sherman@drake.edu

*Flipped instruction is being implemented in an increasing number of mathematics classes but the research base is not yet well developed. Many studies of flipped instruction involve a small number of flipped classes being compared to non-flipped classes, but this methodology fails to account for variations in implementations. To aid in the systematic attention to variation, this article presents a framework for flipped mathematics instruction that identifies key features of the videos assigned as homework as well as features of the in-class activities. The components of the framework are accompanied by proposed quality indicators to further distinguish between flipped implementations that are structurally similar but different in enactment*

Keywords: Flipped Instruction; Lesson Observations; Classroom Research; Frameworks

In recent years there has been an increase in flipped instruction (Smith, 2014). Flipped instruction is characterized by teachers "flipping" the settings in which lecture and homework occur. Instead of being presented with new material in class and then completing homework problems outside of class, students in flipped classes watch videos or read new material outside of class and then complete the problem sets in class (Bergmann & Sams, 2012). Although the flipped instructional model has a single overarching label, there are a wide variety of learning opportunities it can encompass. On one hand, flipped instruction has the potential to be innovative with regard to differentiation (e.g., students can rewatch videos at their own pace and with technology-based supports) and collaboration (e.g., students can work together because of the increased time in class to solve problems). On the other hand, flipped instruction can be non-innovative, such as when students watch lecture videos at home and then work individually to complete procedural exercises in class (de Araujo, Otten, & Birisci, 2017a).

Because of the variation in flipped implementations, this theoretical article presents a research-based framework that moves beyond a general definition of flipped instruction and provides specific analytic tools that allow for meaningful distinctions to be drawn with respect to the learning opportunities afforded. In particular, the framework considers the components of flipped mathematics lessons and the quality indicators for each component.

## Why a Framework is Needed

Although flipped instruction can vary in its implementation (de Araujo, Otten, & Birisci, 2017a), much of the early literature on flipped instruction seems to presume that it is a unified approach that can be directly contrasted with non-flipped instruction. For example, DeSantis, Van Curen, and Putsch (2015) investigated flipped and non-flipped geometry instruction in two high school classes. They found no differences in geometry learning, although students in the flipped class were less satisfied with the unit than students in the non-flipped class. But the flipped lessons in the study involved using the same activities as in the non-flipped lesson, so the implicit assumption was that the act of flipping itself—not the manner of flipping—is of central importance. Moreover, DeSantis and colleagues did not provide substantial detail about the non-flipped instruction, so with regard to the finding of no difference, the basis of comparison is under-specified. A similar example is Clark (2015), who compared his flipped Algebra 1 classes' year-end student outcomes with non-flipped Algebra 1 classes taught by others in the same school. Clark found, like DeSantis and colleagues (2015), no difference between flipped and non-flipped learning measures. Unlike DeSantis et al., however, Clark's students expressed a preference for flipped instruction because of the range of teaching practices it allowed. Again, because of the simplistic and small-scale comparison, the underlying assumption in Clark's study is that it is meaningful to compare flipped instruction with non-flipped instruction of the same content, without substantial regard for the variations that are possible within each category.

The fact that these studies found no difference with regard to content learning, and contradictory results with regard to student attitudes about flipped instruction, is unsurprising. Such results are to be expected if it is true that, as we argue, flipped mathematics instruction can vary widely along important dimensions of instruction, just as non-flipped instruction also varies. Nevertheless, these studies provide a positive first step with regard to empirically documenting the phenomena of flipped instruction in mathematics by showing that, in certain situations, flipping alone does not seem to influence student learning. Yet there are substantial limitations to the accumulated knowledge that can be generated from these studies because there is not yet a structural framework for identifying and interpreting the various features of flipped instruction in the cited studies. With such a framework in hand, patterns could be discerned across grade levels and even across articles that include empirical descriptions of flipped instruction without comparisons to non-flipped instruction (e.g., Talbert, 2014; Zack et al., 2015).

It should be noted that other frameworks for flipped instruction exist, but they serve different functions than ours. For example, Chen and colleagues (2014) introduced the FLIPPED model (Flexible environments, Learner-centered approach, Intentional content, Professional educators, Progressive networking, Engaging and effective learning experiences, Diversified and seamless learning platforms) but theirs is a prescriptive model meant to guide the design and implementation of a flipped course. It is not designed to capture the variability across preexisting flipped courses. Furthermore, it is not specific to mathematics instruction and it functions at the course level whereas our framework is applied at the lesson level. DeLozier and Rhodes (2017) provided a more descriptive framework, parsing the out-of-class lectures from the in-class activities and specifying different types of the latter (quizzes, pair-and-share activities, student presentations). In our work, however, we have seen that not all flipped instructors use strictly lecture videos (de Araujo, Otten, & Birisci, 2017b), and DeLozier and Rhodes do not account for the quality of the enacted in-class activities, nor is their work specific to mathematics. Yet we do incorporate their top-level distinction between out-of-class and in-class lesson phases.

To assess the quality of the flipped lessons in a mathematics-specific manner, we draw on existing instruments for lesson observations but make some key modifications. Observation protocols such as the M-SCAN (Walkowiak et al., 2014) and IQA (Boston & Wolff, 2006) presume a certain form of instruction, with cognitively demanding tasks and student discussions, and thus may not be widely applicable to all forms of mathematics instruction. Moreover, these instruments combine various sub-scores into an overall score for the quality of the lesson, even though there is scholarly disagreement about the ideal nature of some of the sub-scores. In other words, although some aspects are universally regarded as indicators of quality instruction (e.g., coherent development of ideas, effective use of multiple representations), other aspects are not universally accepted (e.g., the value of shared mathematical authority, student discussions) (Munter, Stein, & Smith, 2015). Thus, in order to capture the full range of flipped instruction variability in an even-handed way, it would be wise to separate the universally-accepted indicators of quality from those that simply mark different instructional paradigms. We represent this separation in the in-class component of our framework.

**A Framework for Flipped Mathematics Instruction**

Adapting the overall structure of the Mathematical Task Framework (Stein, Grover, & Henningsen, 1996), our framework (Figure 1) includes the assignment given to students to be completed outside of class (left) and the activity that takes place in class (center), which collectively produce various student outcomes (right). We distinguish between different types of videos (lecture and set-up) and between different interaction formats in class (whole-class and non-whole-class). These distinctions are rooted in our perspective on students' mathematical learning as a sociocultural process through which students are coming to participate actively in a mathematical community (Vygotsky, 1980). This process is an integration of social and individual dimensions and, mathematically, it involves becoming an active participant both in individual meaning-making through thinking or writing, for example, as well as collaborative meaning-making through listening to or interacting with the teacher and peers, for example. In other words, we view the learning of mathematics as not simply the development of an ability to solve certain kinds of problems but the development of the ability to participate in various forms of mathematical discourse and activity, and our framework distinguishes between some of these forms at a structural level. We also attend to the nature of discourse (Staples & Colonis, 2007) and the cognitive demand of tasks (Stein, Grover, & Henningsen, 1996) because these relate to the kinds of activities students have opportunities to participate in.

This framework is based on pilot analyses of flipped classes at several different academic levels and consultation with experts in specific domains (e.g., video design, classroom discourse, mathematical tasks). Due to space limitations, details on the development process are not shared.

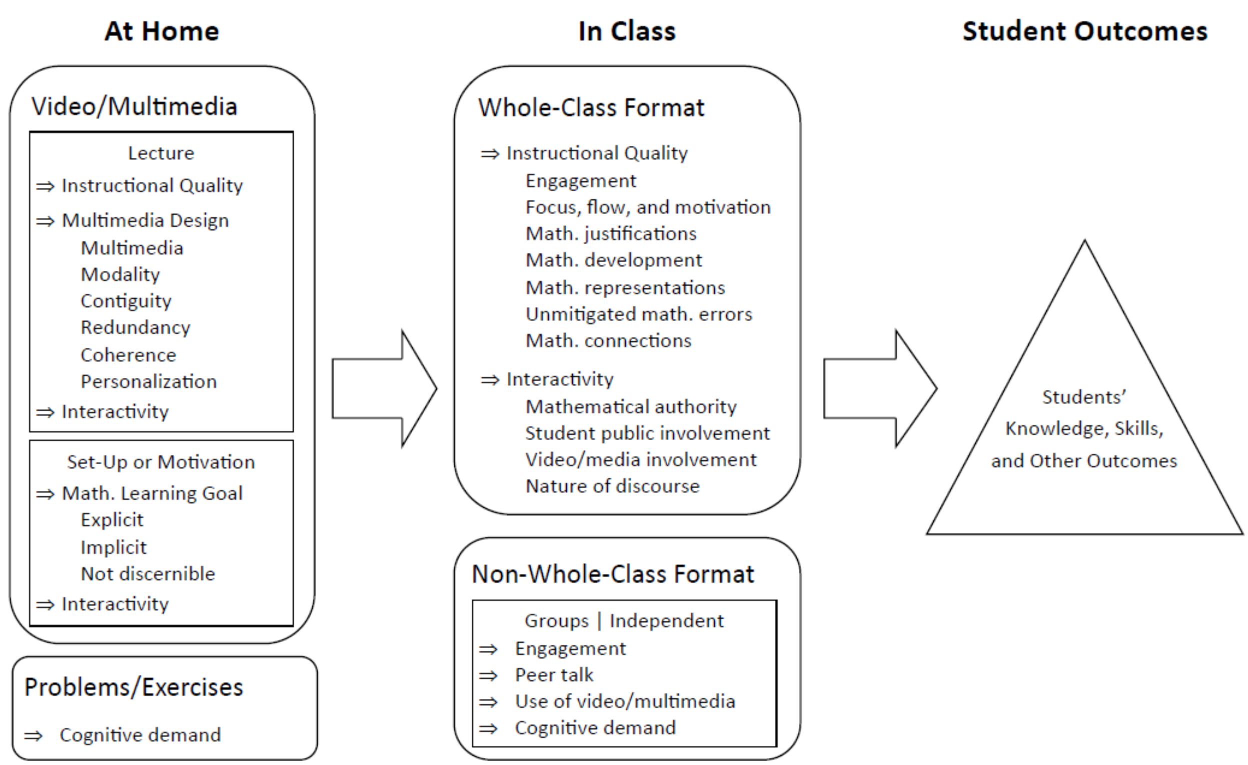


**Figure 1.** A framework for flipped mathematics instruction

Hodges, T.E., Roy, G. J., & Tyminski, A. M. (Eds.). (2018). *Proceedings of the 40th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Greenville, SC: University of South Carolina & Clemson University.

**At-Home Components**

Looking more closely at the *at-home* components, the framework identifies *video or multimedia* homework in which students are assigned multimedia content to watch, read, or listen to. Although this type of homework is the definitive category for flipped lessons, teachers implementing flipped instruction are not bound to it exclusively. Teachers may, on occasion, also assign *problems or exercises* to be completed outside of class, in which case the framework could be used to show that both or either type of homework was assigned. Also, it is possible that a particular lesson has no homework assigned. Non-flipped lessons, by definition, will have only problems/exercises homework or no homework.

**Videos or multimedia.** Although we often use the term 'video' because flipped mathematics homework commonly takes this form, we acknowledge that some teachers might use written text, podcasts, animations, or multimedia resources such as iBooks. For video homework, the framework distinguishes between *lecture* videos, which present expository information or worked examples, and *set-up or motivation* videos, which establish a context or pose a problem that will be addressed in class. Within these two types, the framework then specifies normative quality criteria. By normative quality, we mean that instruction satisfying the criteria is hypothesized to be more predictive of positive student outcomes than instruction that does not satisfy the criteria, which could be empirically tested in future studies.

The normative quality criteria for lecture videos address three aspects: instructional quality, multimedia design, and interactivity. *Instructional quality* captures the clarity and flow of the ideas as well as features specific to mathematics, such as mathematical justifications, multiple mathematical representations, the absence of unmitigated errors, and connections to mathematical content that the students may have already learned or will learn in the future. Most of these particular aspects can be operationalized by drawing on existing instruments in the mathematics education literature. In particular, the clarity of the lecture *foci* relates to research on explicit lesson objectives (Wiggins & McTighe, 2005), and the *coherent flow* of ideas relates to all of the video content being connected to that foci. We can also attend to whether there is an intellectual or "real-world" *motivation* for the foci. With regard to mathematics-specific features, we draw on the Studying Teacher Expertise and Assignment in Mathematics (STEAM) project (Tarr et al., 2016), which provides a lesson observation coding scheme for *mathematical justifications*, the *meaningful development of mathematical ideas*, and the *integration of multiple representations*. In addition to these aspects from STEAM, we also include a dimension from the mathematical quality of instruction (MQI) instrument (Learning Mathematics for Teaching Project, 2011), namely, the *absence of unmitigated mathematical errors*. Finally, we attend to whether the video's particular mathematical foci is explicitly *connected to the mathematical ideas* of prior or subsequent lessons.

These features of instructional quality not only apply to the videos but are also used to characterize whole-class discourse in the in-class portion of lessons (see below). Aspects that are specific to videos, however, are multimedia design and interactivity. For *multimedia design* quality, we drew on Clark and Mayer's (2007) six principles of digital material design. The *multimedia principle* states that graphics (not words alone) help students learn the information in question. The *modality principle* states that the graphics should be described or explained with spoken words rather than on-screen text. The *contiguity principle* states that, if some brief on-screen text is used, it should be placed spatially near the relevant graphics. The *redundancy principle* states that the on-screen text should be minimal and should not be a full transcription of the audio narration. The *coherence principle* states that irrelevant graphics and audio should be excluded. Finally, the *personalization principle* states that audio narration should be in a conversational style and it is preferable for the speaker to be visible.

*Interactivity* captures the extent to which students are expected to be more than passive viewers of the video. Interactivity can be achieved in a variety of ways. For example, Webel, Sheffel, and Conner (2018) studied a teacher whose videos for flipped instruction contained embedded questions that provided the teacher with formative assessment data she could use to inform in-class activities. These embedded questions within the video are one possible form of interactivity. Other forms include requiring students to post a comment or submit a reflection after watching the video or incorporating a virtual manipulative (Moyer-Packenham & Westenskow, 2013) that the students can interact with as part of the homework.

The aspects described above apply to a lecture video, which is the most common type of video in flipped lessons (de Araujo, Otten, & Birisci, 2017b). Our framework, however, also includes *set-up or motivation videos*. Because of a lack of prior research on set-up videos, we drew on practical resources for setting up mathematical work such as the three-act tasks promoted by Meyer (2011) often involving what we call a set-up video that presents a situation and leads to a mathematical question. Our framework distinguishes between different ways in which the set-up videos connect to the mathematical goals of the remainder of the lesson. Specifically, we distinguish between three possibilities: an *explicit connection* to the mathematical goal of the in-class activity that follows (i.e., it is clear from watching the video what the specific mathematical question or problem is, even though its solution is not explicated); an *implicit connection* (i.e., it presents a context that allows for mathematical explorations but it is not clear which mathematical question or problem is intended to be pursued until it is specified later); or *no discernible connection* (i.e., it has a contextual connection or some other association with the in-class activity but does not contain the *mathematical* ideas of the lesson). We do not presume a priori that one of these is more beneficial than the others, but that could be determined empirically through applications of the framework.

**Problems/exercises**. If students are assigned a problem set to complete outside of class, then the framework attends to the cognitive demand (Stein, Grover, & Henningsen, 1996) of those problems. After coding each item for cognitive demand, the assignment overall can be characterized as *low* (nearly all items at a low level), *moderate* (predominantly low but a substantial minority of items at a high level), or *high* (a majority of the items at a high level). Note that if problems are begun in class and students are then expected to complete any unfinished problems at home, this would fall under the in-class component of group or independent work rather than strictly homework.

Moving on, the arrow between the at-home section of the framework and the in-class section represents the fact that what occurs outside of class potentially influences what occurs in class. For example, if students do not watch the video, the lesson will likely play out differently in class than if they had watched the video attentively. Note that, although videos are typically assigned to be watched prior to the in-class portion of the lesson, some flipped lessons may use a video as a way to summarize ideas after the in-class activity. Thus the arrow is not necessarily chronological and the left-to-right order visible in Figure 1 is not absolute.

**In-Class Components**

Moving to the *in-class* components, the framework first distinguishes between the interaction format the teacher employs—a *whole-class format* in which the entire class is expected to be attending to the public discourse whether that be the teacher addressing the entire class or students sharing out or asking questions, and a *non-whole-class format* which includes the expectation that students work together in *groups* or *independently*.

**Whole-class format**. As noted above, this framework differs from others because we separate instructional quality indicators from markers of a particular pedagogical model. For instructional quality we use the same aspects defined above (focus, flow, motivation, mathematical justifications, mathematical development of ideas, multiple mathematical representations, absence of unmitigated mathematical errors, and connections to other mathematical topics) plus an engagement (i.e., on-task) measure to gain a normative sense of the lesson flow and development of mathematical ideas. Additionally, the framework distinguishes between several descriptive aspects of the lesson. These aspects relate to the interactivity of the whole-class segments of the lesson. With regard to *mathematical authority*, we distinguish between teacher/textbook authority and shared/community authority. We also note the extent to which *students are publicly involved* in the whole-class discourse, and if they are, we attend to whether the *nature of discourse* was sharing or collaborative (Staples & Colonis, 2007). Finally, considering flipped lessons in particular, we note whether the *videos are referenced or incorporated* into the discourse (e.g., by reviewing the key ideas from the video homework).

**Group or independent work**. With regard to group or independent work, the framework includes *engagement* (i.e., on-task behavior), *peer talk* (i.e., the extent to which students verbally interact with one another), and the *use of video/multimedia* from the lesson homework. For example, while solving a problem in class, do students use their phones or tablets to view videos and rewatch explanations of worked examples? We also use the well-known construct of *cognitive demand* (Stein, Grover, & Henningsen, 1996) to characterize the written tasks students work on during group or independent work. Because flipped instruction can provide more in-class time for problem solving, it is plausible that teachers may feel more freedom to enact cognitively demanding tasks in their flipped lessons.

### Student Outcomes

The framework does not predetermine the student outcomes of interest because those will largely depend upon the specific context in which the framework is employed, especially with regard to specific mathematical content. However, the framework does acknowledge that several different student outcomes may be of interest in studies of flipped instruction. In addition to measures of students' mathematical knowledge, one may also be interested in their skills with various mathematical practices and their attitudes or affective dispositions toward mathematics. Furthermore, researchers may wish to use multiple measures of a single outcome to increase the robustness of findings. In any event, the specification of these outcomes is somewhat independent of the enactment of lessons through the at-home and in-class phases of the lesson.

## Discussion

In presenting this framework, we hope to provide a structure that can support the accumulation of empirical findings related to flipped instruction, inform the design of future studies, and facilitate connections between research and practice. This framework is needed because empirical research on the effectiveness of flipped instruction must move beyond simple flipped/non-flipped comparisons (e.g., Clark, 2015; DeSantis et al., 2015) toward an investigation of the specific features of flipped instruction and the complex interactions of those features as they relate to student outcomes.

Although the framework is built around structural components of flipped lessons, it also goes further by focusing attention on the quality with which certain structural activities occur. In particular, the framework includes quality criteria for lecture videos, set-up videos, whole-class discourse, and it incorporates the levels of cognitive demand (Stein, Grover, & Henningsen, 1996) for the independent or group tasks. This attention to quality will not only allow researchers to distinguish between different structural implementations of flipped instruction (e.g., those who use lecture videos and independent work versus those who use set-up videos and whole-class discussions) but also between different qualities of implementation within the same structure (e.g., productive whole-class discussions versus unproductive ones). To be clear, we do not mean to imply that the quality criteria specified in this article are the only sets of criteria that could be used, nor do we necessarily think they are ideal. Rather, we have proposed some characteristics that, according to our theoretical perspective, are worth considering, but those with different perspectives would likely identify different characteristics, and ultimately we hope that future research can empirically test the benefits of various implementations of flipped instruction.

With regard to empirical studies, the framework can be valuable in formulating testable hypotheses. For example, one can consider whether video features are more predictive of student learning than in-class features. (We hypothesize that the reverse is true.) Additionally, the framework can be used to aggregate across lessons rather than remain at the scope of an individual lesson. Future research could include multiple observations of a flipped class over time and compile the proportional allotments and the quality indicators for the components of flipped instruction. For instance, a teacher may 80% of the time assign lecture videos that satisfy only some of the quality criteria and may assign the other 20% of the time set-up videos that explicitly connect to the lesson's learning goals, and 50% of the in-class time may be spent solving high cognitive demand tasks in groups while the other 50% of the time is spent having high quality whole-class discussions. In this way, multiple class aggregations could be formed (for flipped and non-flipped classes) and then analyzed with respect to student outcome measures. Such work has potential to contribute to our theoretical understanding of mathematics instruction and to inform the efforts of practitioners who are implementing flipped instruction.

## Conclusion

Many of the expositional and anecdotal works related to flipped instruction are written in ways that suggest flipped instruction is a unified teaching model. Even the few empirical studies that exist involved research designs that presume flipped instruction is a *singular approach* that can be meaningfully compared to non-flipped instruction. We have instead found that wide variation exists in flipped implementation (de Araujo, Otten, & Birisci, 2017b; Otten, Zhao, de Araujo, & Sherman, in press). Although it is possible that flipped instruction frees up class time for rich whole-class discussions and the enactment of cognitively demanding tasks, we know that there are secondary mathematics teachers using flipped instruction in ways that differ little from conventional non-flipped instruction aside from the change in instructional environment. As a field we do not yet know the prevalence of either of these implementations nor do we have compelling empirical evidence of the implications that either has for student learning. It is our hope that this framework will guide feature endeavors to gather this evidence.

## Acknowledgments

This work was supported by the ReSTEM Institute at the University of Missouri and the National Science Foundation (NSF award #1721025). Any opinions, findings, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of NSF.

## References

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Eugene, OR: International Society for Technology in Education.

Boston, M. D., & Wolf, M. K. (2006). *Assessing academic rigor in mathematics instruction: The development of Instructional Quality Assessment Toolkit* (CSE Tech. Rep. No. 672). Los Angeles: University of California, CRESST.

Chen, Y., Wang, Y., Kinshuk, & Chen, N.-S. (2014). Is FLIP enough? Or should we use the FLIPPED model instead? *Computers and Education, 79*, 16–27.

Clark, K. R. (2015). The effects of the flipped model of instruction on student engagement and performance in the secondary mathematics classroom. *Journal of Educators Online, 12*, 91–115.

de Araujo, Z., Otten, S., & Birisci, S. (2017a). Mathematics teachers' motivations for, conceptions of, and experiences with flipped instruction. *Teaching and Teacher Education, 62*, 60–70.

de Araujo, Z., Otten, S., & Birisci, S. (2017b). Conceptualizing "homework" in flipped mathematics classrooms. *Journal of Educational Technology and Society, 20*, 248–260.

DeLozier, S. J., & Rhodes, M. G. (2017). Flipped classrooms: A review of key ideas and recommendations for practice. *Educational Psychology Review, 29*, 141–151.

DeSantis, J., van Curen, R., Putsch, J., & Metzger, J. (2015). Do students learn more from a flip? An exploration of the efficacy of flipped and traditional lessons. *Journal of Interactive Learning Research, 26*(1), 39–63.

Learning Mathematics for Teaching Project. (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education*, *14*, 25–47.

Meyer, D. (2011, May 11th). The three acts of a mathematical story. Retrieved from http://blog.mrmeyer.com/2011/the-three-acts-of-a-mathematical-story/

Munter, C., Stein, M. K., & Smith, M. S. (2015). Is there a common pedagogical core? Examining instructional practices of competing models of mathematics teaching. *NCSM Journal of Mathematics Education Leadership, 16*(2), 3–13.

Otten, S., Zhao, W., de Araujo, Z., & Sherman, M. (in press). Evaluating videos for flipped instruction. *Mathematics Teaching in the Middle School*.

Smith, D. F. (2014, June 12). How flipped classrooms are growing and changing. *Ed Tech Magazine*. Retrieved on October 27, 2015, from http://www.edtechmagazine.com/k12/article/2014/06/how-flipped-classrooms-are-growing-and-changing

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*, 455–488.

Talbert, R. (2014). Inverting the linear algebra classroom. *PRIMUS, 24*, 361–374.

Tarr, J. E., Campbell, P. F., Webel, C., Markworth, K. A., & Gibbons, L. (2016, April). *Examining the impact of elementary mathematics specialists and coaches.* Paper presented at the 2016 National Council of Teachers of Mathematics Research Conference, San Francisco, CA.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Walkowiak, T. A., Berry, R. Q., Meyer, P., Rimm-Kaufman, S. E., & Ottmar, E. R. (2014). Introducing an observational measure of standards-based mathematics teaching practices: Evidence of validity and score reliability. *Educational Studies in Mathematics, 85*, 109–128.

Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.

Zack, L., Fuselier, J., Graham-Squire, A., Lamb, R., & O'Hara, K. (2015). Flipping freshman mathematics. *PRIMUS*, *25*, 803–813.